2021 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW) | 978-1-6654-1012-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/EuroSPW54576.2021.00032

Ontology-Based Scenario Modeling for Cyber Security Exercise

Shao-Fang Wen, Muhammad Mudassar Yamin, Basel Katt
Department of Information Security and Communication Technology
Norwegian University of Science and Technology
Gjøvik, Norway
{*shao-fang.wen, muhammad.m.yamin, basel.katt*}*@ntnu.no*

***Abstract*—The growing demand for cyber security professionals with practical knowledge is boosting the development and conduct of cyber security exercises around the world. Scenarios stand a central position of the exercise, which sets the stage for later action by providing contextual information that the participants will need during the exercise. To manage the increasing numbers of scenario creation in the different contexts, we propose an ontology to model scenarios for cyber security exercises. This ontology identifies aspects of scenario modeling relevant to cyber security that can be used as a means to achieve a defined taxonomy of knowledge items and a standard vocabulary for cyber scenarios. With the semantic framework based on RDF/OWL, this ontology provides a common structure at a semantic level that allows scenarios to be shared and reused across applications and community boundaries. In this paper, we present the design, implementation, and evaluation of the proposed ontology.**

***Index Terms*—cyber security exercise, scenario modeling, ontology**

1. Introduction

Cyber Security Exercise (CSE) is increasingly seen as an important part of cybersecurity training in both the private and public sector [26]. CSE has been identified as an effective technique to stimulate cyber security awareness [16], which provides opportunities and an ultimate learning experience [3] for the students or cyber professionals to improve their skills in protecting and defending information systems in the context of a realistic, true-to-life situation [6]. It helps uncover gaps in organizational security policies, procedures, and resources [12], [21], from which employees of an organization can be provided necessary training and/or tools and policies can be rectified too [34]. The growing demand for cyber security professionals with practical knowledge [38] is boosting the development and conduct of different types of CSE around the world. These include technical exercises, such as Locked Shields, which is conducted under the aegis of the NATO Cooperative Cyber Defence Centre of Excellence (CCDCOE) in Tallinn. In such exercises, participants use digital tools to mitigate and counter cyber incidents. There are also non-technical exercises such as the annual Cyber 9/12 Challenge, an annual desk-based activity, in which students from across the globe compete in developing policy solutions tackling a fictional cyber catastrophe. There are also small-scale exercises conducted inside corporate or government contexts [12].

All the training operations carried out in a CSE revolve around the concept of scenario. Scenarios describe the situation that the exercise is trying to simulate and the logical flow of events that will be happening [18]. A scenario could include the digital infrastructure where all the activities are staged in as well as all the game-like aspects such as targets, vulnerabilities, and rules of engagement [16], [26]. Due to the large creation of CSE nowadays and the increasing complexity of scenarios, how to efficiently deal with scenario information, such as information sharing and scenario inference, is a non-trivial problem. Additionally, because cyber scenarios fundamentally involve various contexts, containing distributed IT systems, complex operational tasks and security concepts (e.g., vulnerabilities), knowledge management of all these aspects is necessary to the retrieval and reuse of scenarios. As interoperability and computability of scenario information in CSE promise a great deal in the global effort for better quality and more efficient cyber security training, yet remain largely unachieved.

To address the aforementioned issues, we use ontologies as a means to achieve a defined taxonomy of knowledge items and a standard (conceptual) vocabulary for defining cyber scenarios to achieve knowledge standardization and sharing. The ontology provides a common structure at a semantic level that allows data to be shared and reused across applications, enterprises, and community boundaries [23]. Semantic Web technologies provide rich constructs to represent information that is not only machine-readable but also machine-understandable, thus facilitating semantic integration and sharing of information from heterogeneous sources. The main benefit of the ontology-based model is the availability of a formal, encoded description of the security knowledge: that is, all the entities, their attributes, and their interrelationships will be defined and represented, described in detail within [23], [30]. To the best of our knowledge, the cyber-scenario ontology presented in this paper is the first ontology work based on the CSE context, which unifies cyber security knowledge with general cyber scenarios about entities and relations. By adopting the ontology, CSE teams can manage scenarios efficiently, share and reuse knowledge and employ the reasoning capability of the semantic web to draw inferences among different scenarios. In addition, users can extend the core level scenario ontology to satisfy the specific requirements of different specific domains in cyber security. The remainder of this paper is organized as follows. In Section 2, we introduce the scientific background of the research. The

DOI 10.1109/EuroSPW54576.2021.00032

design of the ontology is explained in Section 3, while the implementation and evaluation are described in Section 4 and Section 5 respectively. Finally, the discussion and conclusions are presented in Section 6.

2. Scientific Background

This section presents a necessary scientific foundation of the research subject area. The theoretical and practical underlying topics are discussed. The topics include an overview of cyber security exercises and cyber secenarios, and the concepts of ontologies.

2.1. Cyber Security Exercise

An exercise, as described in the ISO Guidelines for Exercises [25], is "a process to train for, assess, practice, and improve performance in an organization". The National Institute of Standards and Technology (NIST) also gives a definition for exercises in the special publication 800-84 (Guide to Test, Training, and Exercise (TTE) Program for IT Plans and Capabilities [18]): "An exercise is a simulation of an emergency designed to validate the viability of one or more aspects of an IT plan." European Cyber Security Organization (ECSO) defines a CSE as "a planned event during which an organization simulates cyber-attacks or information security incidents or other types of disruptions to test the organization's cyber capabilities, from being able to detect a security incident to the ability to respond appropriately and minimize any related impact." [14]. According to Yamin et al. [43], in general term, CSE is "a training exercise that runs attack and/or defense scenarios on virtual and/or physical environments intending to improve the attack and/or defense understandings and skills of the participants."

Cyber security exercises provide opportunities for participants to apply theoretical, hypothetical concepts in a training environment without fear of adversely affecting the "real world" [24]. This kind of training technique is therefore invaluable as they focus on participants' activities; maximize participation; are motivational and give immediacy to the subject matter, something of great benefit to a fast-moving, rapidly changing cyber environment [12]. CSE can be carried out by small, individual entities such as single ministries or private firms or, in the case of large, multinational simulations, exercises can involve a multitude of actors from different areas of the security nexus, such as private corporations, government ministries, utility providers and military units [10]. In a CSE, personnel with roles and responsibilities in a particular cyber security plan meet to validate the content of a plan through discussion of their roles and their responses to emergencies, execution of responses in a simulated operational environment, or other means of validating responses that do not involve using the actual operational environment for deployment of personnel. In other words, CSE encourages participants to use skills, techniques, tools, and policy frameworks they know in a practical, virtual environment in order to be better prepared should a real cyber crisis ensue.

CSEs can be generally divided into two different categories [18], [21]: *discussion-based exercises* and *operation-based exercises*. In a discussion-based exercise, participants are presented with a situation or question-related to the scenario that they are required to discuss and formulate the appropriate response or solution among the participants of roles, responsibilities, coordination, and decision-making. Operation-based exercises normally involve multi-agency participation (real or simulated) and they can focus on one or many geographical areas. This type of exercise is used to practice multiple emergency functions e.g. direction and control, resource management, and communications. Operation-based exercises allow participants to interact within a simulated environment with their roles and responsibilities through an exercise control group that provides prewritten injects and responds to questions and tasks developing out of the exercise.

2.2. Cyber Scenario

CSEs are scenario-driven, such as a power failure in one of the organization's data centers or a cyber-attack causing certain systems to be crashed, with additional situations often being presented during an exercise. Scenarios have been termed a 'story and simulation' approach, in which storylines about how relevant events might unfold in the future are used to parameterize models of cyber-physical and social processes, each consistent with an alternative future [2], [27]. Scenarios use logical implications, assumptions, and forecasts to communicate about a potential future state [4]; it incorporates issues to be resolved, time relations, interactions, and consequences [19], [27]. Scenarios are most powerful when several are used together to present alternative views of the future as seen from the present because humans and organizations can take actions that influence the future [19].

An exercise's scenario is a sequential, narrative account of a hypothetical incident that provides the catalyst for the exercise and is intended to introduce situations that will inspire responses and thus allow demonstration of the exercise objectives [18]. Exercise scenarios may be formulated to tackle worst-case situations or typical security issues that participants are apt to encounter in the real world. For example, an exercise of a network contingency for an organization that is prone to disruptions from external events may consider a scenario involving a significant power outage caused by a natural disaster, such as a hurricane.

In general, scenarios can be described in two formats [31]:

- *Outlined scenarios*: an outlined scenario provides a general summary of the damage and disruption to people, property, and services that have resulted. It sells the idea of the exercise and facilitates information gathering. Examples of outlined scenarios can be found in [7] and [35].
- *Detailed scenarios*: a detailed scenario contains additional information and is much more comprehensive than an outlined scenario. It has several parts that sequentially describe the event's impact on specific services or sections of the organization, along with a timeline for the restoration of key lifeline utilities. Examples of detailed scenarios can be found in [18] and [40].

In the context of CSEs, a scenario defines the training environment as well as real situations that will inspire

responses that help participants achieve exercise objectives [43]. Simply put, the CSE scenario is a story that sets out the fictive conditions and events of the exercise. It is typically based on a cyber security problem and its background. Cyber scenario integration into exercises focuses on the inclusion of the cyberspace domain as an integral element of the operational environment, not as a separate event or scenario. A narrative scenario is documented and typically distributed to participants via handouts or an oral presentation before the exercise or at the beginning of the exercise event. Scenario documentation usually includes a brief scenario background, summaries, orders, and a storyline designed to provide participants a sense of the world/local situation, ensuring the representative operational context supports training objectives [37]. The scenario itself portrays the events that will occur during the conduct of the event. These events will also become a part of the Master Scenario Events List (MSEL) (typically used in operations-based exercises), which serves as the script for the execution of the exercise; it includes the ordering of injects, time of execution, and the expected reactions from the training audience [18].

In summary, the scenario has two important functions:

- It sets the mood for the exercise, captures the participant's attention, and motivates them to continue.
- It also sets the stage for later action by providing information that the participants will need during the exercise.

2.3. Ontologies

One of the most accepted definitions of ontologies in the field of computer science is the one given by Gruber [20], who defines an ontology as "an explicit formal specification of a conceptualization". This can be further elaborated that an ontology is a formal description of the relevant concepts and relationships in an area of interest, simplifying and abstracting the view of the world for some purpose [42]. An ontology is a technology that provides a way to exchange semantic information between people and systems. It provides a common vocabulary and depicts all the concepts and inter-concept relations in a formal logic representation. An ontology is a graph whose nodes represent the concepts or objects of a domain, and the edges indicate relationships between concepts. Usually, this graph is structured around a hierarchical "backbone" similar to the class/subclass relationship in object-oriented programming. Due to the formalization, it can be represented and to some degree interpreted by machines and enables the formal analysis of the domain. This allows an automated or computer-aided extraction and aggregation of knowledge from different sources and possibly in different formats [20].

Ontologies play an important role in achieving interoperability across organizations and on the Semantic Web [15] because they aim to capture domain knowledge and their role is to create semantics explicitly in a generic way, providing the basis for agreement within a domain Semantic Web technologies provide representation languages, such as Resource Description Framework (RDF) and Web Ontology Language (OWL), to represent the semantics of an entity as a set of things or concepts rather than strings of words. Ontologies are now central to many applications such as scientific knowledge portals, information management, and integration systems, electronic commerce, and web services. The main areas, in which ontological modeling is applied, include communication and knowledge sharing, logic inference and reasoning, and knowledge base. By analyzing and extending several types of research [29], [41], we can identify and summarize the reasons for and benefits of developing and using ontologies in knowledge modeling.

- Ontologies share a common understanding of structured information among people or software agents.
- Ontologies make domain knowledge reusable.
- Ontologies enable the interoperability among models or specific domain vocabularies.
- Ontologies allow and simplify the communication among humans, computational systems, and between humans and systems.
- Ontologies have the expressive power for acquiring context from diverse and heterogeneous sources.

3. An Ontology-Based Scenario Model for Cyber Security Exercise

In this section, we present the modeling of the scenario ontology for the cyber security exercises. We aim to structure cyber scenario elements and the corresponding security knowledge. To develop this ontology, we first reviewed the literature in the domains of cyber security exercises and cyber security knowledge modeling to identify an inventory of the elements and relationships of a cyber scenario model. Subsequently we sliced the ontology into three submodels: *Scenario Information Model*, *Scenario Operation Model*, and *Security Knowledge Model*. Figure 1 illustrates the interrelation between the three models at the semantic level. The scenario information model describes basic information about the exercise scenario, which implies security knowledge in the security knowledge model. The scenario operation model describes the concepts of injects, representative systems, and teams that are critical for the operationalization of the scenario as well as of the implied security knowledge. Sections 3.1 – 3.3 below describe the detailed design of the three models in the ontology.

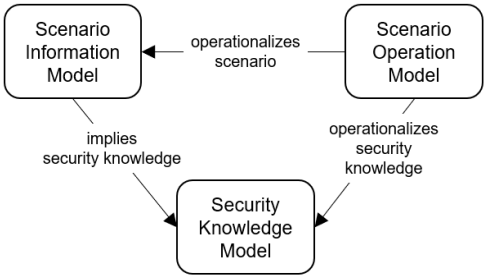

Figure 1. A semantic view of the ontology design

3.1. Scenario Information Model

Most stakeholders of the exercise, from exercise directors, controllers, and simulators, to the players are involved in one way or another with the scenario itself. The scenario repositories, therefore, are always focused on storing the information perspective first. The information model describes the context and main attributes of scenarios (see Figure 2). The key concepts in the scenario information model are described in the following.

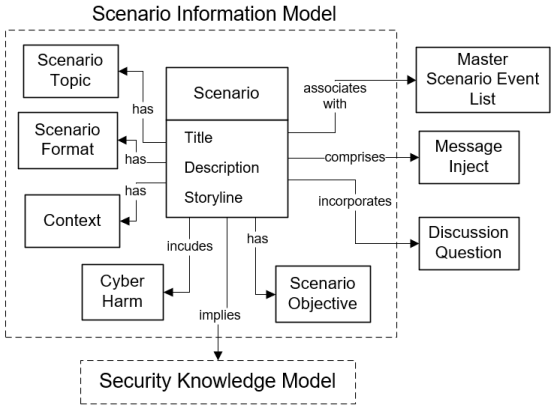

Figure 2. Scenario information model

A *Scenario*, as mentioned in Section 2.2, can be defined as an existing or potential cyber incident. The class of Scenario stands as the kernel of the ontology, which provides linkages (relationships) with the security model and two concepts: *Master Scenario Event List* and *Inject* in the scenario operation model. The class of Scenario is modeled with the following basic attributes: *title*, *description*, and *storyline*. A storyline includes a series of events described that make up a narrative. Comparing to scenario descriptions, a storyline gives meaning to a broader vision, which tells a story through a timeline in the overall scenario (e.g., pre-event, post-incident, the morning of the event, etc.).

A *Scenario Format* refers to the description granularity of the scenario narratives, which can be distinguished by outlined and detailed, as mentioned in Section 2.2.

A *Scenario Topic* is a combination of words and phrases used to name a scenario's key concepts for search and retrieval purposes. Examples of topics for a single scenario could be the combination of "phishing email", "ransomware", and "antivirus".

A *Context* refers to the circumstances or conditions that form the setting for a scenario. *Context* has the following sub-classes:

- *Sector*: It refers to the sectors of the society and economy, where there is an impact on the scenarios. For example, government, transport, nuclear, etc.
- *Domain*: It refers to the application domain of the scenario, e.g., IoT, cloud, telecommunication, etc.
- *Cyber System*: A cyber system refers to any combination of facilities (hardware and software), digital content, and communications integrated to provide cyber-service that are used (or mentioned) in the scenario.
- *Actor*: An actor in a scenario is a participant engaging in an action or process directly or indirectly (e.g., attackers, hackers, threat actors, security personnel, users, etc.)
- *Social environment*: A social environment refers to the social and cultural setting in which actors live or in which something happens or develops in a scenario. Examples of social environments are organizations, teams, social events, etc.

A Cyber Harm is the damaging consequence demonstrated in the cyber scenario. It corresponds to the consequences of an attack or a threat. Unlike the abstract descriptions of security knowledge in the *Security Knowledge Model*, the cyber harm is contextualized, that is, described concretely to the context of the scenario. By modeling this concept specifically in the ontology, users can easily search for the failure points in the real world and draw on relevant security knowledge.

3.2. Scenario Operation Model

The *Scenario Operational Model* describes how organizational resources (people, processes, and systems) can be arranged or configured to perform the scenario. There are three major inter-related parts: (1) the representative system inventory, which consists of a collection of cyber system components and the configuration or rules that govern the deployment of these components, (2) the sequencing events/injects, which describe the detailed activities of the scenario and expected actions from the participants, and (3) the teams (see Figure 3).

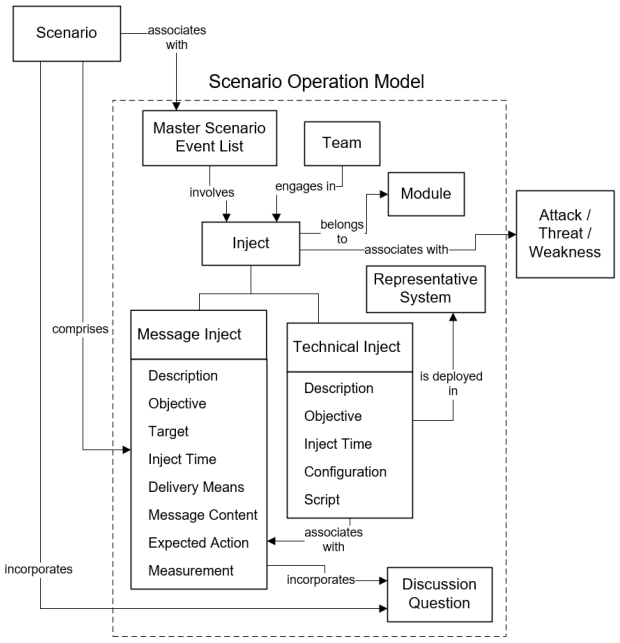

Figure 3. Scenario operation model

A *Master Scenario Event List* (MSEL) is a chronologically sequenced outline of the simulated events and key injects (with messages or computer scripts) for the operationalization of the scenario. MSEL plays an important role in operation-based exercises since it's used to ensure timely and organized execution of the exercise [18],

while in discussion-based exercises, a formal MSEL is not necessary. An *Inject* is also known as an implementer or injector, that will be demonstrated to participants during an exercise. In practice, an inject is usually associated with an attack event or a threat that exploits vulnerabilities presented in the scenario context. Injects are designed to elicit a specific response from participants, such as triggering action and a responsible decision that is in line with security countermeasures or organizational incident response plan. There are two types of injects: *Message Inject* and *Technical Inject*.

- *Message Inject*: A message inject is a pre-scripted message that can be delivered in different means, for example, email, letters, radio, and telephone calls. Each message inject contains information designed to supplement the scenario, prompt additional actions, or trigger discussions toward predefined questions. They expand on the outline of the scenario storyline or the key events portrayed in the MSEL. The attributes of an inject include *description*, *target*, *inject time*, *delivery means*, *message content*, *expected action*, and *measures of performance*. An Expected Action represents the management/administration's desired responses or actions to the questions or messages proposed during the delivery of injects. The Measure of Performance describes the criteria for assessing participants' response to the message inject. It is important to be aware that, in discussion-based exercises, especially in the detailed-type scenario, message injects are directly linked with the scenario itself, while in operation-based exercises, injects are commonly contained in MSEL.
- *Technical Inject*: A technical inject is a goal-oriented action in the operation-based exercise, which is used to emulate attacker behavior or network traffic according to the context of the scenario. In association with message injects, technical injects are often carried out with various degrees of automation with less (or without) human involvement, such as attack vectors, vulnerability injectors, or real-use traffic simulators (gray traffic), and can be delivered to the representative system by employing scripts or system configuration deployment. A typical technical inject involves the following attributes: *description*, *inject time*, *script*, and *configuration*.

A *Representative System* is defined as a group of information systems that can represent (or simulate) the affected cyber objects in the scenario. When practicing scenarios, it is preferable to work with representations of cyber systems (also known as a testbed or infrastructure orchestrators), since it is neither safe nor legitimate to execute a scenario on an operational cyber system. The class of representation system contains the following subclasses:

- *Equipment*: An equipment is a set of physical resources serving to equip the scenario operation, e.g., personal computers, mobile devices, servers, etc.
- *Node*: A node is used to define a virtual machine (VM) in the virtual environment.
- *Network*: The term network in this model refers to a group of elements used to represent the network information of cyber-system components in a network topology. The sub-classes include *IP address*, *router*, *subnet*, and *port*
- *Installed Software*: Software can be installed in equipment or nodes and is classified into two subclasses: *application software* and *system software*.
- *Service*: It refers to a software functionality or a set of software functionalities (such as the retrieval of specified information or the execution of a set of operations). Examples are email services, printing services, etc.
- *Credential*: Credentials refer to the verification of identity or tools for authentication. They may be part of a (digital) certificate or other authentication processes that help confirm a user's identity concerning a network address or other systems.
- *Content*: Content is any content that exists in the form of digital data. Forms of digital content include information that is digitally transmitted or contained in computer files (e.g., images, videos, webpages, data, etc.). Cryptocurrencies are also a type of digital content.
- *Configuration*: A configuration refers to how components of representative systems are arranged, and how their options are set. System configurations are usually deployed with configured scripts.

A *Module* represents a specific training topic in the overall scenario. Events and injects are typically included within modules, i.e.,

$$Scenarios > Modules > Events/Injects$$

The scenario section is usually divided up into distinct, sequenced modules, which are always based on exercise objectives and scenario requirements. More advanced exercises can contain this level of depth while designing the scenario. Examples of modules are phishing emails, data recovery, social media takeover, etc.

A *Team* is a placeholder for the roles of exercise participants. The concept of *Team* is usually applied in the operation-based exercises, in which different colors are assigned to participants to identify their roles. For example, the red team represents the 'attackers' (or 'offenders'), while the blue team acts as the defender.

A *Discussion Question* is designed to address specific problems or issues that link back to the scenario objectives. Dependent on the scenario format, discussion questions can be incorporated into scenario itself or an inject. Responses to the discussion questions are the focus of the scenario/inject, and reviewing them provides the basis for performance evaluation.

3.3. Security Knowledge Model

To train the trainees in learning attack techniques and knowledge of various types of vulnerabilities, and train on the identification of damage situation and response actions, it is necessary to associate the scenario with security domain knowledge. The *security knowledge model* describes the security-concept modeling and relationships with other entities in the ontology (see Figure 4). The

security knowledge model is developed to serve as a core knowledge base for security-knowledge management and better knowledge reuse. We have surveyed and reviewed existing cybersecurity standards and ontologies to incorporate the most commonly used taxonomies in the ontology and to provide a common understanding of the cybersecurity domain. The modeled security concepts and the corresponding taxonomy are described in the following.

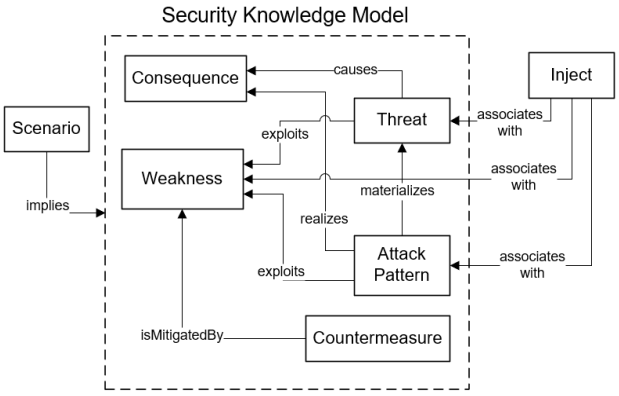

Figure 4. Security knowledge model

An *Attack Pattern* is a description of the common attributes and approaches employed by adversaries to exploit known weaknesses in cyber-enabled capabilities [5]. While attack trees [33] provide a holistic view of the potential attacks facing a particular piece of software, attack patterns provide actionable detail on specific types of common attacks potentially affecting entire classes of information systems. To model attack patterns, we utilize Common Attack Pattern Enumeration and Classification (CAPEC) [8], which organizes attack patterns hierarchically based on mechanisms that are frequently employed when exploiting a vulnerability. According to the analysis of cyber security attack taxonomy [11], CAPEC was identified as outperforming all the other taxonomies.

A *Threat* is a potential negative action/event/condition/circumstance facilitated by a vulnerability that results in an unwanted impact on a computer system or application. Many threat taxonomies already exist in the literature, defined by different security organisms. In our ontology, we adopt the European Union Agency for Cybersecurity (ENISA) Threat Taxonomy [13], which summarizes cyber threats that have been accessed by collecting publicly available information. The most important aspect of the threat taxonomy is that it is open to the addition of new threats to the hierarchical tree without modifying its inherent structure. The layout conceived by ENISA is presented as a table with three levels of classification.

A *Weakness* is a type of mistake that, in proper conditions, could contribute to the introduction of vulnerabilities within that product [9], regardless of whether it occurs in implementation, design, or other phases of a product lifecycle. Instead of modeling *Vulnerability*, we take *Weakness* in the ontology to categorize similar types of vulnerabilities that aim to efficiently facilitate the identification, mitigation, and prevention effort. Security weaknesses, from a broad view, cover social (e.g., organizations and processes), human (e.g., behavior), and technical aspects (e.g., systems and technologies). In this regard, we mainly apply the taxonomy scheme provided by the Software Engineering Institution (SEI) [6], which categorizes cyber security risks as those due to the failures from four perspectives: (1) actions of people, (2) internal processes, (3) external events and (4) systems and technologies. We keep the detailed classification (i.e., the subclasses) of the first three perspectives, which contain human and organizational factors. However, for the system and technologies perspective, Common Weakness Enumeration (CWE) [28] provides a more comprehensive classification of weaknesses in hardware and software that can either be a faulty configuration in the hardware or vulnerabilities present in the software. We then model the CWE security weakness taxonomy under the class of *System and Technologies*.

A *Countermeasure* is an action, device, procedure, or technique that reduces a threat, a vulnerability, or an attack by preventing or mitigating it, by minimizing the harm it can cause, or by discovering and reporting it so that corrective action can be taken. To properly classify cyber security countermeasures, we adopt NIST Cyber Security Framework [36]. It provides a common organizing structure for multiple approaches to cybersecurity by assembling standards, guidelines, and practices that are working effectively today. The taxonomy is structured with five high-level functions: *identify*, *protect*, *detect*,*respond*, and *recover*.

A *Consequence* is a result or effect, typically one that is unwelcome or unpleasant. In the context of cyber security, the consequence is the harm caused to an exploited organization by a cyber-attack. Agrafiotis et al. [1] present a taxonomy of organizational cyber-harm which should help researchers and practitioners to consider the full range of harms that might result from cyber-attacks. The main cyber-harm they include are (1) *physical or digital harm*, (2) *economic harm*, (3) *psychological harm*, (4) *reputational harm*, and (5) *social and societal harm*.

Figure 5 shows the complete ontology model including the interrelationships of the models.

4. Implementation of the Ontology

To implement the designed ontology, we used Protégé [39], a free, open-source platform that provides a suite of tools to construct domain models and knowledge-based applications with ontologies. A part of the class structure is depicted in Figure 6. Some of the subclasses are omitted in this figure to simplify the presentation. We have implemented the full scope of the security taxonomies in Protégé described in Section 3.3 (see Figure 6 (c)). The relationships between the classes are maintained with object properties. Some of these object properties are shown in Figure 7.

Figure 8 demonstrates the individual creation in Protégé with modeling object properties and data properties of a scenario, named '`S004: Ransomware infection`'. This scenario was constructed in a detailed scenario format with five message injects. This scenario describes a phishing attack that employs ransomware via email, encountered in an institutional financial department. Figure 9 depicts one of the injects '`S004-Inject1`' in the scenario '`S004`'. This inject was maintained with critical data properties

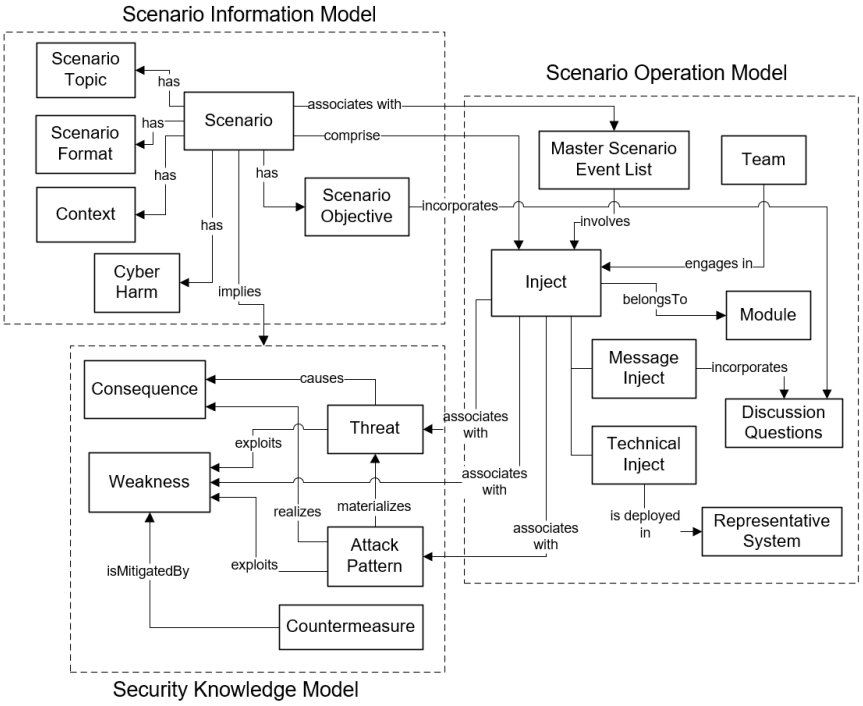

Figure 5. The proposed scenario ontology for cyber security exercises

(a) (b) (c)

Figure 6. Configuration of classes and subclasses in Protégé: (a) Scenario information model (b) Scenario operation model (c) Security knowledge model

that could be used for discussion-based exercises, including the message content and the discussion questions. The related security knowledge elements (attacks and threats) were also made associated with this inject. All the classes and corresponding relationships (object properties and data properties) are described using OWL, which leads to an XML representation of scenarios for a platform-independent, Internet-based interaction of domain experts with our scenario ontology. Figure 10 depicts the snippet of OWL RDF/XML showing an individual (‘S004’) in the scenario class.

Figure 7. Part of the object properties

5. Evaluation of the Ontology

In this section, we evaluate the ontology by running queries using the SPARQL protocol [22] on the ontology. These SPARQL queries answer specific use cases of the

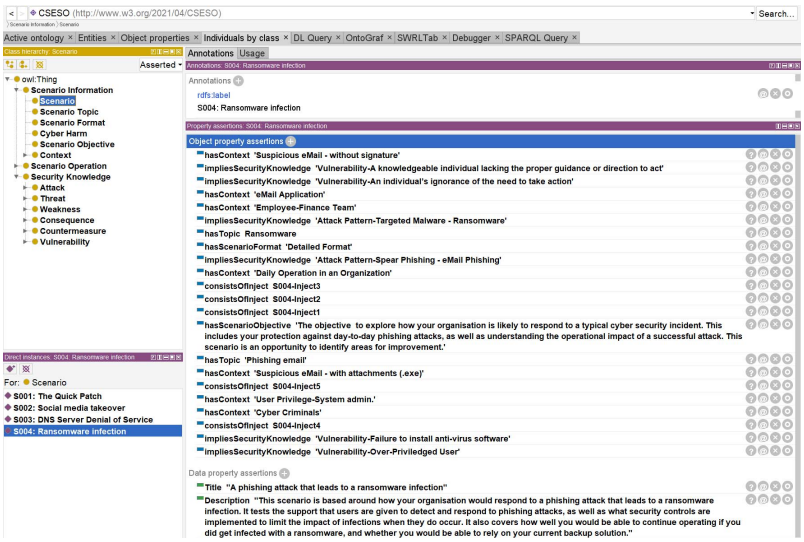

Figure 8. Defining individuals in Protégé: A scenario

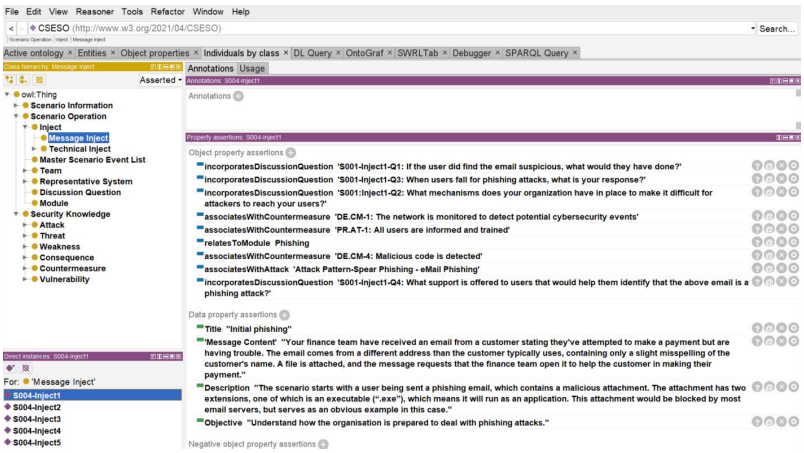

Figure 9. Defining individuals in Protégé: An inject

```
<!-- http://www.w3.org/2021/04/CSESO#S004 -->

<owl:NamedIndividual rdf:about="http://www.w3.org/2021/04/CSESO#S004">
    <rdf:type rdf:resource="http://www.w3.org/2021/04/CSESO#Scenario"/>
    <CSESO:hasContext rdf:resource="http://www.w3.org/2021/04/CSESO#Attacker_Cyber_Criminals"/>
    <CSESO:hasContext rdf:resource="http://www.w3.org/2021/04/CSESO#Digital_Content_email_wo_Signature"/>
    <CSESO:hasContext rdf:resource="http://www.w3.org/2021/04/CSESO#EmailAttachment"/>
    <CSESO:hasContext rdf:resource="http://www.w3.org/2021/04/CSESO#Employee-Financial_Team"/>
    <CSESO:hasContext rdf:resource="http://www.w3.org/2021/04/CSESO#SocialContext01"/>
    <CSESO:hasContext rdf:resource="http://www.w3.org/2021/04/CSESO#User_Privliege_System_Administrator"/>
    <CSESO:hasContext rdf:resource="http://www.w3.org/2021/04/CSESO#eMail_Application"/>
    <CSESO:hasScenarioFormat rdf:resource="http://www.w3.org/2021/04/CSESO#DetailedFormat"/>
    <CSESO:hasScenarioObjective rdf:resource="http://www.w3.org/2021/04/CSESO#SecenarioObjective_S004"/>
    <CSESO:hasTopic rdf:resource="http://www.w3.org/2021/04/CSESO#PhishingEmail"/>
    <CSESO:hasTopic rdf:resource="http://www.w3.org/2021/04/CSESO#Ransomware"/>
    <CSESO:impliesSecurityKnowledge rdf:resource="http://www.w3.org/2021/04/CSESO#CAPEC-163"/>
    <CSESO:impliesSecurityKnowledge rdf:resource="http://www.w3.org/2021/04/CSESO#CAPEC-542"/>
    <CSESO:impliesSecurityKnowledge rdf:resource="http://www.w3.org/2021/04/CSESO#STF-SW-01"/>
    <CSESO:impliesSecurityKnowledge rdf:resource="http://www.w3.org/2021/04/CSESO#STF-SW-02"/>
    <CSESO:impliesSecurityKnowledge rdf:resource="http://www.w3.org/2021/04/CSESO#V1.3.2-1"/>
    <CSESO:impliesSecurityKnowledge rdf:resource="http://www.w3.org/2021/04/CSESO#V1.3.3-1"/>
    <CSESO:Description>This scenario is based around how your organisation would respond to a phishing attack that lea
    current backup solution.</CSESO:Description>
    <CSESO:Objective>The objective  to explore how your organisation is likely to respond to a typical cyber security
    <CSESO:Title>A phishing attack that leads to a ransomware infection</CSESO:Title>
    <rdfs:label>S004: Ransomware infection</rdfs:label>
</owl:NamedIndividual>
```

Figure 10. Snippet of OWL RDF/XML showing an individual in the scenario class

competency questions (CQs). This method of evaluation is considered a very effective evaluation technique to test the adaptability and consistency of an ontology [32]. If the SPARQL queries can extract individuals as a response, it signifies that the CQs have succeeded in covering the defined objectives of the ontology. Therefore, three exemplary CQs were developed considering the interoperability among different models.

CQ 1. What is the context of a given scenario 'S004'?

CQ 2. List the inject information (including title, module, message content, and discussion questions) related to scenario 'S004'.

CQ 3. What are the implied attack patterns in the scenario 'S004' and what are the corresponding weakness?

The corresponding SPARQL statement for each evaluated CQ and the execution result in the Protégé editor are depicted in Figure 11 - 13.

```
SELECT  DISTINCT ?Context ?ContextElement
WHERE {
   ?Scenario onto:hasContext ?ContextElement.
   ?ContextElement rdf:type ?Context.
    FILTER (regex(str(?Scenario), "S004") &&
           ?Context != owl:NamedIndividual)
}
ORDER BY ?Context
```

Context	ContextElement
Actor-Attacker	Cyber Criminals
Actor-User	Employee-Finance Team
Digital Content	Suspicious eMail - with attachments (.exe)
Digital Content	Suspicious eMail - without signature
Privilege	User Privilege-System admin.
Social Enviroment	Daily Operation in an Organization
Software	eMail Application

Figure 11. The SPARQL statement and the execution result of CQ 1

```
SELECT ?MessageInject ?Title ?Module ?MessageContent
       ?DiscussionQuestion
WHERE {
   ?Scenario onto:Comprises ?MessageInject.
   ?MessageInject onto:Title ?Title.
   ?MessageInject onto:belongsToModule ?Module.
   ?MessageInject onto:InjectMessage ?MessageContent.
   ?MessageInject onto:incorporates ?DiscussionQuestion.
   FILTER regex(str(?Scenario), "S004").
}
ORDER BY ?Inject
```

Inject	Title	Module	MessageContent	DiscussionQuestion
S004-Inject1	"Initial phishing"	Phishing	"Your finance team have receive	"What mechanisms does your or
S004-Inject1	"Initial phishing"	Phishing	"Your finance team have receive	"What support is offered to usen
S004-Inject1	"Initial phishing"	Phishing	"Your finance team have receive	"If the user did find the email sus
S004-Inject1	"Initial phishing"	Phishing	"Your finance team have receive	"When users fall for phishing atb
S004-Inject2	"Virus infection"	Antivirus	"The email recipient from the fina	"Is this user likely to have an up
S004-Inject2	"Virus infection"	Antivirus	"The email recipient from the fina	"How would the antivirus have r
S004-Inject2	"Virus infection"	Antivirus	"The email recipient from the fina	"What measures would be in pla
S004-Inject3	"Over-Prvileged User"	User Permission	"When the attached file is openec	"Do you have any method of blo
S004-Inject3	"Over-Prvileged User"	User Permission	"When the attached file is openec	"Would the user (and therefore t
S004-Inject3	"Over-Prvileged User"	User Permission	"When the attached file is openec	"What permissions do your usen
S004-Inject4	"Ransomware and backups	Ransomware	"After several hours, users have s	"How long could you continue to
S004-Inject4	"Ransomware and backups	Ransomware	"After several hours, users have s	"Are your backups separate to y

Figure 12. The SPARQL statement and the execution result of CQ 2

6. Discussion and Conclusion

Ontologies have acquired a crucial role in enhancing the value of knowledge management in the domain of cyber security which facilitates reuse, sharing, and management of security knowledge efficiently and effectively. In this paper, we propose a scenario ontology for cyber security exercises, which is organized in three sub-models: the scenario information model, the scenario operation model, and the security knowledge model. The main contribution of the paper consists in (1) identifying aspects

```
SELECT ?AttackPattern ?ResourceA ?Weakness ?ResourceW
WHERE {
   ?Scenario onto:impliesSecurityKnowledge ?AttackPattern.
   ?AttackPattern rdfs:seeAlso ?ResourceA.
   ?AttackPattern onto:exploits ?Weakness.
   OPTIONAL { ?Weakness rdfs:seeAlso ?ResourceW }
   FILTER regex(str(?Scenario), "S004").
}
```

AttackPattern	ResourceA
CAPEC-542: Targeted Malware	"https://capec.mitre.org/data/definitions/542.html"
CAPEC-542: Targeted Malware	"https://capec.mitre.org/data/definitions/542.html"
CAPEC-163: Spear Phishing	"https://capec.mitre.org/data/definitions/163.html"
CAPEC-163: Spear Phishing	"https://capec.mitre.org/data/definitions/163.html"

Weakness	ResourceW
CWE-250: Execution with	"https://cwe.mitre.org/data/definitions/269.html"
CWE-358: Improperly Impl	"https://cwe.mitre.org/data/definitions/358.html"
Inaction of People: A know	
Inaction of People: An indi	

Figure 13. The SPARQL statement and the execution result of CQ 3

of scenario modeling relevant to cyber security exercises, (2) implementing them in a semantic framework based on RDF/OWL. The ontology provides definitions of general terms that can be used for constructing a cyber scenario. The scenario model is complemented by commonly used cyber-security taxonomies, which provides a comprehensive view of security knowledge associated with the scenario. Moreover, this ontology, when populated, makes up a scenario library with a collection of instances representing the scenarios. The ontology could then be used as a basis to translate the scenario instance into XML to enhance interoperability or into any other language that is deemed appropriate for various knowledge manipulation processes. Scenario developers can use our proposed ontology to build a repository of scenarios and scenario elements. The repository could contain past scenarios, i.e. previously security incidents, simulated scenarios, and scenario components suitable for reuse. Furthermore, it is advantageous to archive scenarios from previous exercises together with results for future reference. An advantage of using this type of repository approach is that it makes it easier to work with several types of exercises, which could correspond to actual contexts, among multiple scenarios. Such systematic methods in managing scenarios using the ontology help cyber exercise planners to construct scenarios efficiently, and to run them in the simulation environment.

We hope this paper contributes to strengthening research in the area and fostering debate on the concrete role and value provided by ontologies in this domain. We have not yet formally examined the ontology in regard with operation-based-exercise scenarios, for example, in the context of cyber range. In the future, the ontology will be extended and updated as further implementation is rolled out: more stakeholders and tasks will be identified. The authors will involve domain experts to evaluate the contents of the ontology with more case studies. The future work should also address a comprehensive, systematic review of the use of ontologies in cyber security exercises that provides the complete picture of the state of the art in the area that has only been succinctly depicted here. That would provide the required roadmap to a more systematic exploration of the potential of ontologies in the cyber security exercises.

References

[1] Agrafiotis, I., et al. 2018. "A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding how they propagate". Journal of Cybersecurity, volume 4, issue 1, pages tyy006.

[2] Alcamo, J. and T. Ribeiro. 2001. "Scenarios as tools for international environmental assessment". volume 5. European Environment Agency.

[3] Augustine, T. and R.C. Dodge. 2006. "Cyber defense exercise: meeting learning objectives thru competition".

[4] Avery, A. 2020. "Cybersecurity Scenario Modeling: Imagining the Black Swans for Digital Infrastructures Risk Management".

[5] CAPEC, "CAPEC Glossary-Attack Pattern"; Available from: https://capec.mitre.org/about/glossary.htmlAttackPattern. (Accessed on Feb 3, 2021)

[6] Cebula, J.L. and L.R. Young. 2010. "A taxonomy of operational cyber security risks", Carnegie-Mellon Univ Pittsburgh Pa Software Engineering Inst. pages.

[7] Center for Internet Security, "Tabletop Exercises: Six Scenarios to Help Prepare Your Cybersecurity Team "; Available from: https://www.cisecurity.org/wp-content/uploads/2018/10/Six-tabletop-exercises-FINAL.pdf. (Accessed on Feb. 02, 2021)

[8] Common Attack Pattern Enumeration and Classification (CAPEC), "CAPEC VIEW: Mechanisms of Attack"; Available from: https://capec.mitre.org/data/definitions/1000.html. (Accessed on Feb 2, 2021)

[9] CWE, "CWE Glossary-Weakness"; Available from: https://cwe.mitre.org/documents/glossary/index.htmlWeakness. (Accessed on June 3, 2019)

[10] Department of Homeland Security. 2006. "Cyber Storm Exercise Report".

[11] Derbyshire, R., et al. 2018. "An analysis of cyber security attack taxonomies". in 2018 IEEE European Symposium on Security and Privacy Workshops (EuroSPW). IEEE.

[12] Dewar, R.S. 2018. "Cybersecurity and cyberdefense exercises", ETH Zurich.

[13] ENISA, "Threat Taxonomy"; Available from: https://www.enisa.europa.eu/topics/threat-risk-management/threats-and-trends/enisa-threat-landscape/threat-taxonomy/view. (Accessed on Feb. 2, 2021)

[14] European Cyber Security Organisation (ECSO), "Understanding Cyber Ranges: From Hype to Reality"; Available from: https://www.ecs-org.eu/documents/uploads/understanding-cyber-ranges-from-hype-to-reality.pdf. (Accessed on Jan. 23, 2021)

[15] Fensel, D., et al. 2001. "Ontologies and electronic commerce". IEEE Intelligent Systems, volume 16, issue 1, pages 8-14.

[16] Furtună, A., V.-V. Patriciu, and I. Bica. 2010. "A structured approach for implementing cyber security exercises". in 2010 8th International Conference on Communications. IEEE.

[17] Garb, Y., S. Pulver, and S.D. VanDeveer. 2008. "Scenarios in society, society in scenarios: toward a social scientific analysis of storyline-driven environmental modeling". Environmental Research Letters, volume 3, issue 4, pages 045015.

[18] Grance, T., et al. 2006. "Guide to test, training, and exercise programs for IT plans and capabilities". volume, issue, pages.

[19] Gray, P. and A. Hovav. 1999. "Using scenarios to understand the frontiers of IS". Information Systems Frontiers, volume 1, issue 1, pages 15-24.

[20] Gruber, T.R. 1993. "A translation approach to portable ontology specifications". Knowledge acquisition, volume 5, issue 2, pages 199-220.

[21] Gurnani, R., K. Pandey, and S.K. Rai. 2014. "A scalable model for implementing Cyber Security Exercises". in 2014 International Conference on Computing for Sustainable Global Development (INDIACom). IEEE.

[22] Harris, S., A. Seaborne, and E.J.W.C.r. Prud'hommeaux. 2013. "SPARQL 1.1 query language". volume 21, issue 10, pages 778.

[23] Hendler, J., O. Lassila, and T. Berners-Lee. 2001. "The semantic web". Scientific American, volume 284, issue 5, pages 34-43.

[24] Hoffman, L.J., et al. 2005. "Exploring a national cybersecurity exercise for universities". IEEE Security Privacy, volume 3, issue 5, pages 27-33.

[25] ISO. 2013. "ISO 22398:2013 Societal security — Guidelines for exercises". pages.

[26] Karjalainen, M., T. Kokkonen, and S. Puuska. 2019. "Pedagogical aspects of cyber security exercises". in 2019 IEEE European Symposium on Security and Privacy Workshops (EuroSPW). IEEE.

[27] Kim, Y.G. and S. Cha. 2012. "Threat scenario-based security risk analysis using use case modeling in information systems". Security and Communication Networks, volume 5, issue 3, pages 293-300.

[28] MITRE, "Common Weakness Enumeration (CWE) "; Available from: https://cwe.mitre.org/index.html. (Accessed on Feb. 02, 2021)

[29] Noy, N.F. and D.L. McGuinness. 2001. "Ontology development 101: A guide to creating your first ontology".

[30] Obrst, L., P. Chase, and R. Markeloff. 2012. "Developing an Ontology of the Cyber Security Domain". in STIDS. Citeseer.

[31] Planning, Mass Evacuation. 2008. "Directors's Guideline for Civil Defence Emergency Management Groups, Wyd". Ministry of Civil Defence Emergency Management, Wellington.

[32] Raad, J. and C. Cruz. 2015. "A survey on ontology evaluation methods". in Proceedings of the International Conference on Knowledge Engineering and Ontology Development, part of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management.

[33] Saini, V., Q. Duan, and V. Paruchuri. 2008. "Threat modeling using attack trees". Journal of Computing Sciences in Colleges, volume 23, issue 4, pages 124-131.

[34] Samejima, M. and H. Yajima. 2012. "IT risk management framework for business continuity by change analysis of information system". in 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE.

[35] Scarfone, K.A., T. Grance, and K. Masone. 2008. "Sp 800-61 rev. 1. computer security incident handling guide", National Institute of Standards Technology.

[36] Sedgewick, A. 2014. "Framework for improving critical infrastructure cybersecurity, version 1.0".

[37] Staff, U.J. 2012. "Joint Training Manual for the Armed Forces of the United States (CJCSM 3500.03 D)". Washington, DC: Joint Chiefs of Staff, volume, issue, pages.

[38] Topham, L., et al. 2016. "Cyber security teaching and learning laboratories: A survey". Information Security, volume 35, issue 1, pages 51.

[39] Tudorache, T., et al. 2013. "WebProtégé: A collaborative ontology editor and knowledge acquisition tool for the web". Semantic web, volume 4, issue 1, pages 89-99.

[40] U.S. Department of Homeland Security, "Cyber Tabletop Exercise for the Healthcare Industry "; Available from: https://www.hsdl.org/?abstractdid=789781. (Accessed on Jan. 30, 2021)

[41] Uschold, M. and M. Gruninger. 1996. "Ontologies: Principles, methods and applications". The knowledge engineering review, volume 11, issue 2, pages 93-136.

[42] Wand, Y., V.C. Storey, and R. Weber. 1999. "An ontological analysis of the relationship construct in conceptual modeling". ACM Transactions on Database Systems (TODS), volume 24, issue 4, pages 494-528.

[43] Yamin, M.M., B. Katt, and V. Gkioulos. 2020. "Cyber ranges and security testbeds: Scenarios, functions, tools and architecture". Computers Security, volume 88, pages 101636.